

# 基于动态加权差分主成分分析的工业锅炉故障诊断

王文标, 张谦谦

(大连海事大学 船舶电气工程学院, 辽宁 大连 116026)

**摘要:**针对工业锅炉的动态特性与多模态特性带来的故障检测问题,提出一种动态加权差分主成分分析法(DWDPCA)。首先,建立合理的时间窗描述系统的时序特性;然后,对时间窗中的样本寻找其空间上的第一近邻和第一近邻的近邻集,使用加权差分方法将数据转化为单模态结构;最后,利用处理后的数据建立PCA模型进行故障检测。通过在某实际工业锅炉中的应用表明,DWDPCA方法可解决动态时序问题和多模态数据中心漂移问题,显著提高故障检测的精度。

**关键词:**工业锅炉;动态特性;多模态特性;主成分分析;故障检测

中图分类号:TP273 文献标识码:A DOI:10.16146/j.cnki.rndlge.2023.11.022

[引用本文格式]王文标,张谦谦.基于动态加权差分主成分分析的工业锅炉故障诊断[J].热能动力工程,2023,38(11):185-190.  
WANG Wen-biao, ZHANG Qian-qian. Fault diagnosis of industrial boiler based on dynamic weighted difference principal component analysis [J]. Journal of Engineering for Thermal Energy and Power, 2023, 38(11): 185-190.

## Fault Diagnosis of Industrial Boiler based on Dynamic Weighted Difference Principal Component Analysis

WANG Wen-biao, ZHANG Qian-qian

(College of Marine Electrical Engineering, Dalian Maritime University, Dalian, China, Post Code: 116026)

**Abstract:** A novel dynamic weighted differential principal component analysis (DWDPCA) was proposed to address the challenges of detecting faults in industrial boilers with dynamic and multimodal characteristics. Firstly, a reasonable time window was established to capture the time sequence characteristics of the system; then, the first nearest neighbor and its neighbor set in space were searched for the samples within the time window. The data was transformed into a single-modal structure using a weighted differential method; finally, the processed data established a PCA model for accurate fault detection. The application of the DWDPCA method in an industrial boiler, which has shown promising results in solving the dynamic time sequence and center drift problems associated with multimodal data, has significantly improved fault detection accuracy.

**Key words:** industrial boiler, dynamic characteristics, multimodal characteristics, principal component analysis (PCA), fault detection

## 引言

工业锅炉是一种重要的能源转换设备,其长期

处于高温、高压环境中,一旦发生故障将严重危害生命财产安全,因此对锅炉进行运行监控,确保及时、准确地检测到故障具有重要意义<sup>[1-4]</sup>。工业锅炉具有过程变量多、耦合性强等特点,其燃烧过程复杂难

收稿日期:2023-02-09; 修订日期:2023-03-10

基金项目:国家自然科学基金(52071047,62073054)

Fund-supported Project: National Natural Science Foundation of China(52071047,62073054)

作者简介:王文标(1982-),男,大连海事大学讲师。

以建立准确的数学模型对其进行故障检测。基于数据驱动的故障检测方法通过分析和挖掘数据之间的特征关系建立监控模型,即可实现对系统的故障检测。特别是随着分布式控制系统的广泛应用,大量工业过程数据得以采集和保存,为基于数据驱动的故障诊断方法创造了条件<sup>[5-6]</sup>。

以主成分分析(Principal Component Analysis, PCA)为代表的驱动法采用  $T^2$  和 SPE 两个统计量进行运行监控<sup>[7-9]</sup>。但  $T^2$  和 SPE 统计量使用的前提是数据为单模态结构,并且样本之间是相互独立的。工业锅炉虽然是一种典型的单模态连续控制系统,但是当负荷发生变化时,会体现出多模态特性。同时锅炉也是一个动态系统,样本间存在明显的时序特性。因此使用 PCA 方法对锅炉进行故障检测,检测效果并不理想。

针对过程数据存在的多模态问题,He 等人<sup>[10]</sup>提出 k 近邻方法(Fault Detection Using the k Nearest Neighbor Rule, FD-kNN),计算样本与近邻的距离,将累计近邻距离作为监控指标,有效降低数据多模态特性对故障检测的不利影响。但当各模态数据方差差异显著不同时,故障检测性能明显降低。郭金玉等人<sup>[11]</sup>提出一种差分预处理算法,通过寻找样本的最近邻并进行差分处理,使数据聚集在坐标附近,但该方法同样存在对方差差异明显的检测效果不佳的问题。冯立伟等人<sup>[12]</sup>提出基于标准距离度量下的 k 近邻故障检测方法,通过构造标准距离,解决数据多中心和方差差异明显问题。但该方法未考虑样本疏密程度对特征指标的贡献差异,使用该方法进行故障检测,检测效果不佳。Ma 等人<sup>[13]</sup>提出局部近邻标准化方法,通过寻找样本的近邻并使用近邻样本的均值和方差对其进行标准化处理,消除了样本的多模态特性。但当故障样本处于多模态之间时,故障样本的近邻来自多个模态,经该方法处理后无法将故障样本与正常样本分离,降低了故障检测精度。

针对系统的动态问题,Hui 等人<sup>[14]</sup>通过构造时间序列的增广矩阵,消除了样本的时序特性对故障检测的不利影响。但该方法增加了矩阵维度,降低了计算效率。苗爱敏等人<sup>[15]</sup>提出基于时序扩展的邻域保持嵌入算法,通过寻找样本时间上的近邻点,获取样本时序变化特性,但该方法故障检测效果易受近邻数的影响。

针对具有动态特性和多模态特性的锅炉系统,本文提出一种基于动态加权差分主成分分析法(Dynamic Weighted Differential Principal Component Analysis, DWDPCA)的故障检测方法。首先在时间方向上设置合理的时间窗,然后对时间窗中的样本寻找其空间上的第一近邻和第一近邻的近邻集进行加权差分处理,预处理后的数据服从独立同分布且为单模态结构,最后采用 PCA 方法进行故障检测。

## 1 基于动态加权差分主成分分析的故障检测

### 1.1 主成分分析法

原始数据集为矩阵  $X_0 = (X_{0ij})_{n \times m}$ ,其中  $n$  为样本数, $m$  为变量数,对数据进行标准化处理得到标准化矩阵  $X$ ,建立 PCA 模型,如式(1)和式(2)所示:

$$X_i = X_p + E = T_i P_i^T + E \quad (1)$$

$$T_i = X_i P_i \quad (2)$$

式中: $P$ —负载矩阵; $T$ —得分矩阵; $E$ —残差矩阵。

利用 Hotelling's  $T^2$  和 SPE 统计量作为监控指标,当  $T^2$  统计量低于控制限  $T_\alpha$  时,系统正常运行,否则发生故障,即:

$$T_i^2 = t_i \Lambda^{-1} t_i^T = X_i P \Lambda^{-1} P^T X_i^T < T_\alpha \quad (3)$$

式中: $t_i$ — $T$  矩阵的第  $i$  行; $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ ,  $\lambda_i$ —协方差矩阵的特征值。

$T_\alpha$  计算公式为:

$$T_\alpha = \frac{\alpha(n+1)(n-1)}{n(n-\alpha)} F_\alpha(\alpha, n-\alpha) \quad (4)$$

式中: $\alpha$ —检验水平; $F_\alpha(a, n-a)$ —第一自由度为  $\alpha$ 、第二自由度为  $n-a$  的  $F$  分布。

当 SPE 统计量低于控制限  $Q_\alpha$  时,系统正常运行,否则发生故障,即:

$$\text{SPE} = X^T (I - P P^T)^T (I - P P^T) X < Q_\alpha \quad (5)$$

$Q_\alpha$  计算公式为:

$$Q_\alpha = \theta_1 \left( \frac{c_\alpha h_0 \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right)^{1/h_0} \quad (6)$$

式中: $\theta_i = \sum_{j=k+1}^m \lambda_j^i (i = 1, 2, 3)$ ,  $h_0 = 1 - 2\theta_1 \theta_3 / 3\theta_1^2$ ,  $c_\alpha$ —标准正态分布的置信极限。

### 1.2 DWDPCA 方法

由于锅炉过程数据具有动态和多模态特性,使用传统 PCA 方法进行故障诊断不满足  $T^2$  和 SPE 统

计量数据独立且服从单一模态的前提条件,因此检测效果不佳。为了提高 PCA 方法对锅炉系统的故障检测效果,采用 DWDPCA 法,在时间方向上设置合理的时间窗,并对时间窗中的样本进行加权差分处理,再使用 PCA 方法进行故障检测。DWDPCA 方法考虑了样本的时序特性和空间结构,可有效提高 PCA 方法在具有动态和多模态特性系统中的故障检测性能。具体方法如下:

首先对采集到的样本  $x_t$ , 根据时间序列对每个样本进行扩展,通过设置长度为  $N$  的时间窗,将样本  $x_t$  扩展为样本集  $S(x_t) = \{x_1, x_2 \dots x_t \dots x_N\}$ , 扩展后的样本集能更好地反映系统的动态变化特征。然后对时间窗中的每个样本寻找其空间上的第一近邻  $x_f$  和第一近邻的前  $K$  近邻集  $S(x_f^{(x)}) = \{x_1^{(x)}, x_2^{(x)} \dots x_j^{(x)} \dots x_K^{(x)}\}$ , 使用公式(7)对时间窗中的样本进行加权差分处理,消除数据的多模态特性,得到新的样本  $\bar{x}$ 。

$$\bar{x} = \frac{1}{N} \sum_{t=1}^N (x_f - \sum_{j=1}^K w_j^{(t)} m_j^{(t)}) \quad (7)$$

式中:  $m_j = \text{mean}[x_1^{(f_x)}, x_2^{(f_x)} \dots x_j^{(f_x)}]$ —第一近邻样本  $x_f$  的前  $j$  个近邻样本的均值。权值  $w_j^{(t)}$  ( $j=1, 2 \dots k$ ) 由公式(8)确定:

$$w_j^{(t)} = \frac{1}{d(x_f - x_j^{(f_x)})} \bigg/ \frac{1}{\sum_{s=1}^K d(x_t - x_s^{(f_x)})} \quad (8)$$

式中:  $d(x_t - x_j^{(f_x)})$ —第一近邻  $x_f$  与其第  $j$  近邻  $x_j^{(f_x)}$  的欧式距离。二者之间距离越小,两样本属于同一模态的概率越大,所用到的本模态的信息越多,权值越大;相反,二者之间距离越大,几乎使用不到本模态的信息,权值越小。权值满足式(9):

$$\sum_{s=1}^K w_s = 1 \quad (9)$$

最后,对经加权差分处理后的数据使用 PCA 方法进行故障检测。

### 1.3 DWDPCA 故障检测步骤

使用 DWDPCA 方法对锅炉系统进行故障检测,主要分为两步:离线建模和在线诊断,如图 1 所示。具体步骤如下:

离线建模:

(1) 采集锅炉正常工况不同模态下的历史数据,组成训练集  $X$ ;

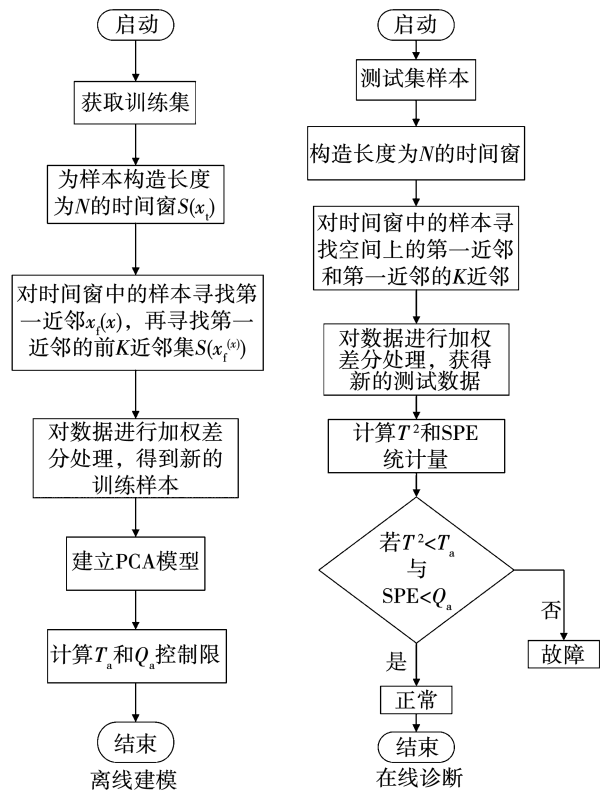


图 1 故障检测流程图

Fig. 1 Flow chart of fault detection

(2) 对训练集中的每个样本构造长为  $N$  的时间窗,对时间窗中的每个样本寻找空间上的第一近邻和第一近邻的前  $K$  近邻集;

(3) 使用公式(7)对数据进行加权差分处理,得到处理后的样本  $\bar{x}$ ,组成新的训练集  $\bar{X}$ ;

(4) 构建 PCA 模型并计算其统计量阈值  $T_\alpha$  和  $Q_\alpha$ ;

在线诊断:

(1) 获取锅炉在线数据,组成测试集  $X^*$ ;

(2) 对测试集中的样本构造长为  $N$  的时间窗,对时间窗中的每个样本寻找空间上的第一近邻和第一近邻的前  $K$  近邻集;

(3) 使用公式(7)对数据进行加权差分处理,得到处理后的样本  $\bar{x}^*$ ,组成新的测试集  $\bar{X}^*$ ;

(4) 为测试集中的每个样本计算其  $T^2$  和 SPE 统计量;

(5) 判断样本的  $T^2$  和 SPE 统计量是否超过其统计量阈值,若超过则锅炉系统发生故障,否则锅炉正常运行。

## 2 工业锅炉运行故障检测

以某 58 MW 供热锅炉为研究对象,采集锅炉在不同负荷下的稳态运行数据,采样时间为 2020 年 11 月 - 2021 年 2 月,采样间隔为 5 min。在锅炉实际运行过程中,由于数据记录错误、停工和设备维修等原因,会造成数据存在异常和错误,因此需要剔除存在明显错误的的数据。经预处理后得到 1 500 组训练集数据,225 组测试集数据,其中 1 - 205 为锅炉正常运行时的数据,206 - 225 为不同状态下人为干预的故障数据。主要监控变量如下:炉膛出口烟气温、炉膛出口温度、省煤器出口烟气温、出水温度、进水温度、炉膛压力、出水流量、引风机转速、鼓风机转速、炉排转速、给煤机转速和烟气含氧量,并使用采样时间作为时间序列变量。

任取锅炉的两个监控变量作为变量  $x_1$  和  $x_2$ ,分别绘制原始数据和经动态加权差分处理后的数据散点图及高斯分布图,如图 2 和图 3 所示,其中  $x_1$  为给煤转速, $x_2$  为烟气含氧量。由图 2 和图 3 可知,原始数据方差差异明显且不服从高斯分布,经动态加权差分方法处理后数据为服从高斯分布的单模态结构。

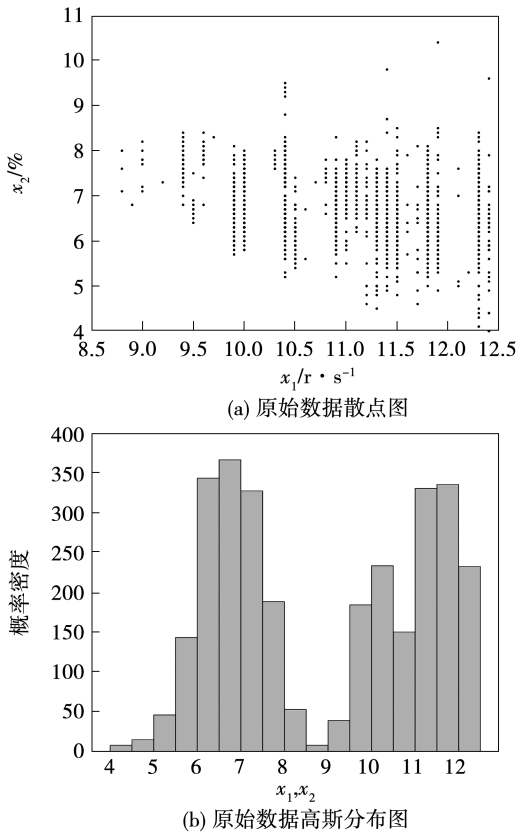


图 2 原始数据散点及高斯分布图

Fig. 2 Scatter and Gaussian distributions of raw data

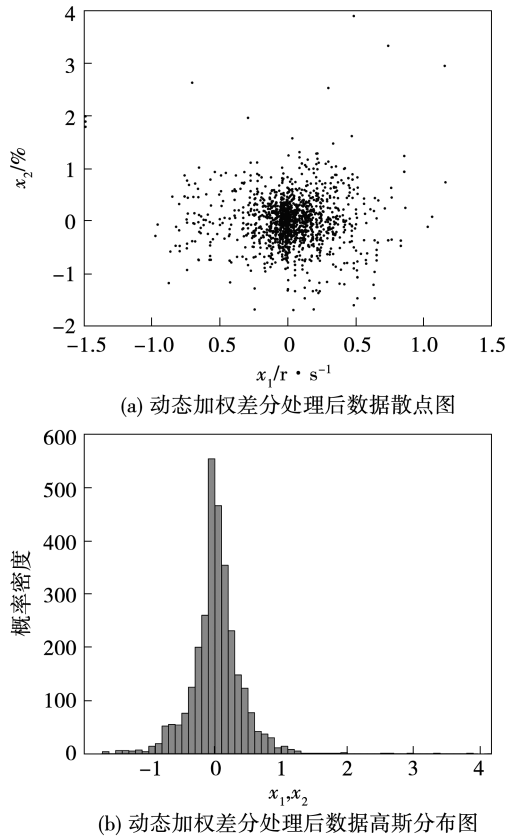


图 3 动态加权差分处理后数据散点及高斯分布图  
Fig. 3 Data scatter and Gaussian distributions after dynamic weighted difference processing

### 2.1 鼓风机故障对比分析

以模拟工业锅炉运行过程中鼓风机故障为例,分别采用 PCA、加权差分 PCA 和 DWDP-PCA 方法进行故障检测,累计贡献率均为 85%,置信度为 99%,检测结果如图 4 所示。由图 4 的 PCA 故障检测图可以看出,使用 PCA 方法无法检测到故障,因此 PCA 方法无法处理具有动态特性和多模态特性的数据,因此故障检效果不佳;加权差分主成分方法将数据转化为服从高斯分布的单模态结构,使得检测效果有所提升,但未解决样本时序问题,因此仍存在部分故障无法检测出来的情况;DWDP-PCA 方法使用时间窗表征样本点之间的自相关性,并将数据转化为单模态结构,使得检测性能进一步提高,可检测到全部故障,因此使用 DWDP-PCA 方法检测效果较为突出。

DWDP-PCA 方法中时间窗长度  $N$  和第一近邻的近邻数  $K$  的选取需要满足经处理后样本之间相互独立且为单模态结构。参数  $N$  的选取将会影响数据动态特性的提取,若  $N$  过大则会造成计算冗余, $N$  过小则当前时间窗无法完整描述数据的时序特性。

同样,若  $K$  过大,则会造成第一近邻与其部分近邻点的欧式距离很大,可用的邻域信息几乎为 0,若  $K$  过小时数据转化为单模态结构将会受到影响。因此,可使用交叉验证的方法来确定参数的大小,本实验中选取  $N$  为 10, $K$  为 5。

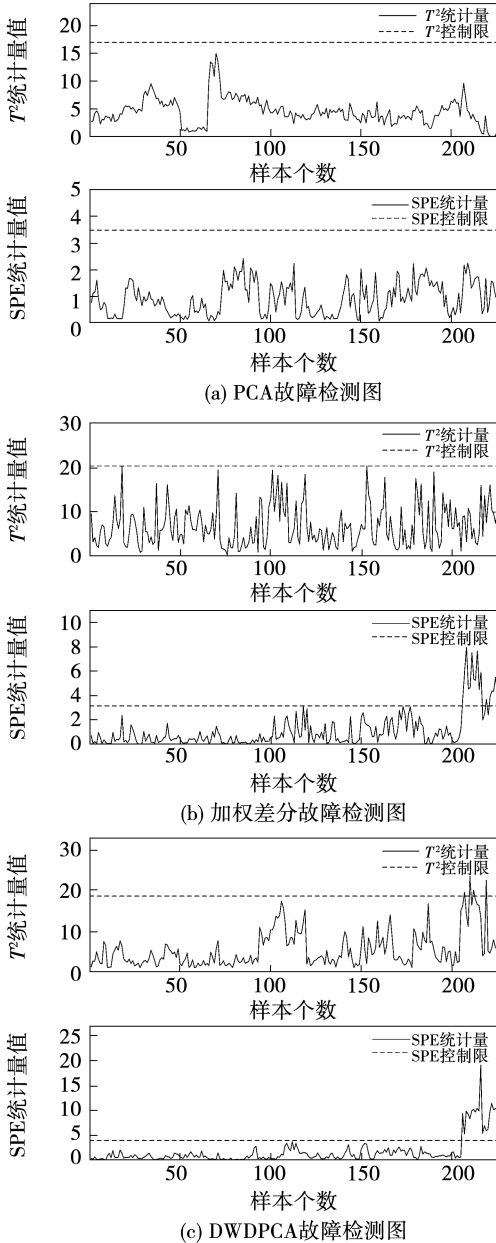


图 4 鼓风机故障检测图  
Fig. 4 Blower fault detection diagram

### 2.2 传感器故障对比分析

在工业锅炉运行过程中,由于恶劣环境影响,传感器易发生故障,通过人工经验很难发现。模拟引风机传感器发生故障,使用 3 种方法进行故障检测,检测结果如图 5 所示。

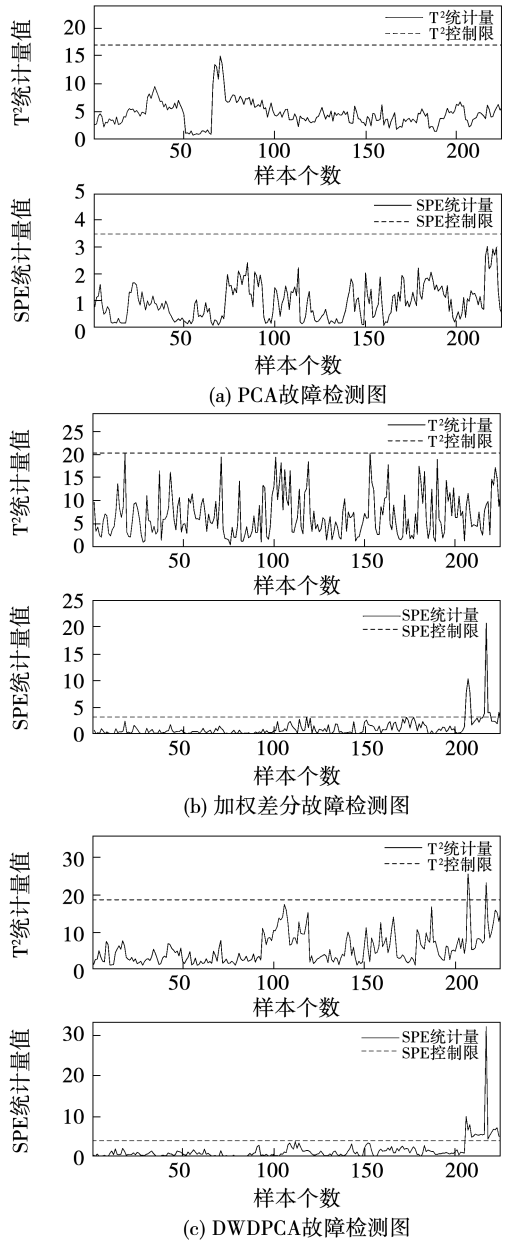


图 5 传感器故障检测图

Fig. 5 Sensor fault detection diagram

由于此次故障只对引风机转速进行了故障模拟,其他变量均保持不变,且故障模拟幅度较小,增加了检测难度。PCA 方法并未检测到故障,这是由于数据具有动态特性和多模态特性,使用该方法计算得到的控制限指标过高,无法检测到故障。加权差分方法可检测 206 ~ 208, 216 ~ 219 和 224 组数据为故障数据。DWDP-PCA 方法可检测到 205 ~ 225 全部故障数据。因此 DWDP-PCA 方法可准确检测出锅炉的运行状态,当运行过程出现异常或故障时,能够及时给出报警。

### 3 结 论

针对工业锅炉中存在的动态特性和多模态特性问题,提出了 DWDPCA 故障检测方法,并将该方法应用于某实际工业锅炉进行故障检测,得出如下主要结论:

(1) 使用 DWDPCA 方法可消除样本的时序特性,并将数据转化为服从高斯分布的单模态结构,消除了动态和多模态特性对故障检测的不利影响。

(2) DWDPCA 方法相比于 PCA 方法和加权差分主成分分析方法具有更高的故障检测率,并通过实验证明了该方法的有效性和优越性。

### 参考文献:

[1] 许裕栗,张 静,李 柠,等.基于数据挖掘的锅炉在线运行状态监测[J].热能动力工程,2019,34(2):82-87,115.  
XU Yu-li, ZHANG Jing, LI Ning, et al. Online operational state monitoring of boiler based on data mining[J]. Journal of Engineering for Thermal Energy and Power, 2019, 34(2): 82-87, 115.

[2] ZHANG S, ZHAO C. Stationarity test and Bayesian monitoring strategy for fault detection in nonlinear multimode processes[J]. Chemometrics and Intelligent Laboratory Systems, 2017, 168: 45-61.

[3] 王文标,田志远,汪思源,等.交叉分段 PCA 在锅炉故障诊断中的应用[J].信息与控制,2020,49(4):507-512.  
WANG Wen-biao, TIAN Zhi-yuan, WANG Si-yuan, et al. Optimal operation method based on cross and piecewise PCA for industrial boilers[J]. Information and Control, 2020, 49(4): 507-512.

[4] 高鹤元,甘辉兵,郑 卓,等.粒子群优化神经网络在船舶辅锅炉故障诊断中的应用[J].计算机应用与软件,2020,37(8):137-141,148.  
GAO He-yuan, GAN Hui-bing, ZHENG Zhuo, et al. Application of particle swarm optimization neural network in fault diagnosis of marine auxiliary boiler [J]. Computer Applications and Software, 2020, 37(8): 137-141, 148.

[5] 洪昌少,黄 俊,关应元,等. BP/RBF 神经网络与模糊规则耦合的电站锅炉燃烧控制[J].热能动力工程,2021,36(4):142-148.  
HONG Chang-shao, HUANG Jun, GUAN Ying-yuan, et al. Combustion control of power station boiler by coupling BP/RBF neural network and fuzzy rules [J]. Journal of Engineering for Thermal Energy and Power, 2021, 36(4): 142-148.

[6] 牛鹏坤,洪 辉,王炜哲.基于改进遗传算法的电站锅炉效率优化[J].热能动力工程,2020,35(3):111-115.  
NIU Peng-kun, HONG Hui, WANG Wei-zhe. Optimization of boiler combustion efficiency based on improved genetic algorithm [J]. Journal of Engineering for Thermal Energy and Power, 2020, 35(3): 111-115.

[7] 郭金玉,王 哲,李 元.基于核熵独立成分分析的故障检测方法[J].化工学报,2022,73(8):3647-3658.  
GUO Jin-yu, WANG Zhe, LI Yuan. Fault detection method based on kernel entropy independent component analysis [J]. CIESC Journal, 2022, 73(8): 3647-3658.

[8] 李 元,姚宗禹.基于邻域保持嵌入的主多项式非线性过程故障检测[J].上海交通大学学报,2021,55(8):1001-1008.  
LI Yuan, YAO Zong-yu. Principal polynomial nonlinear process fault detection based on neighborhood preserving embedding [J]. Journal of Shanghai Jiao Tong University, 2021, 55(8): 1001-1008.

[9] 郭金玉,刘玉超,李 元.一种基于改进局部熵 PCA 的工业过程故障检测方法[J].高校化学工程学报,2019,33(4):922-932.  
GUO Jin-yu, LIU Yu-chao, LI Yuan. A fault detection method based on improved local entropy PCA for industrial processes[J]. Journal of Chemical Engineering of Chinese Universities, 2019, 33(4): 922-932.

[10] HE Q P, WANG J. Fault detection using the K-nearest neighbor rule for semiconductor manufacturing processes [J]. IEEE Transactions on Semiconductor Manufacturing, 2007, 20(4): 345-354.

[11] 郭金玉,仲璐璐,李 元.基于统计差分 LPP 的多模态间歇过程故障检测[J].计算机应用研究,2019,36(1):123-126.  
GUO Jin-yu, ZHONG Lu-lu, LI Yuan. Fault detection of multimode batch process based on statistics difference LPP [J]. Application Research of Computers, 2019, 36(1): 123-126.

[12] 冯立伟,张 成,李 元,等.基于标准距离 k 近邻的多模态过程故障检测策略[J].控制理论与应用,2019,36(4):553-560.  
FENG Li-wei, ZHANG Cheng, LI Yuan, et al. Fault detection strategy of standard-distance-based k nearest neighbor rule in multimode processes [J]. Control Theory & Applications, 2019, 36(4): 553-560.

[13] MA H, HU Y, SHI H. A novel local neighborhood standardization strategy and its application in fault detection of multimode processes[J]. Chemometrics & Intelligent Laboratory Systems, 2012, 118: 287-300.

[14] HUI Y Y, ZHAO X Q. Multiway dynamic nonlinear global neighborhood preserving embedding method for monitoring batch process [J]. Measurement and Control, 2020, 53(5/6): 994-1006.

[15] 苗爱敏,葛志强,宋执环,等.基于时序扩展的邻域保持嵌入算法及其在故障检测中的应用[J].华东理工大学学报(自然科学版),2014,40(2):218-224.  
MIAO Ai-min, GE Zhi-qiang, SONG Zhi-huan, et al. Neighborhood preserving embedding based on temporal extension and its application in fault detection [J]. Journal of East China University of Science and Technology (Natural Science Edition), 2014, 40(2): 218-224.